

PREDICTING KNOT OR CATENANE TYPE OF SITE-SPECIFIC RECOMBINATION PRODUCTS

DOROTHY BUCK AND ERICA FLAPAN

1. SUMMARY

Site-specific recombination on supercoiled circular DNA yields a variety of knotted or catenated products. We develop a model of this process, and give extensive experimental evidence that the assumptions of our model are reasonable. We then characterize all possible knot or catenane products that arise from the most common substrates. We apply our model to tightly prescribe the knot or catenane type of previously uncharacterized data.

Keywords Site-specific recombination, DNA knots, serine recombinases, tyrosine recombinases, DNA topology ^a

2. INTRODUCTION

Since their discovery in the late 1960s, DNA knots and catenanes (*aka* links) have been implicated in a number of cellular processes (see^{1,2} and references therein). In particular, they arise during replication and recombination, and as the products of enzyme actions, notably with topoisomerases, recombinases and transposases.^{2,3} The variety of DNA knots and catenanes observed has made biologically separating and distinguishing these molecules a critical issue.

Experimentally, DNA knots and catenanes can be resolved either via electron microscopy or electrophoretic migration.⁴⁻⁶ Electron microscopy can definitively determine the precise knot or catenane type; however this process can be both difficult – particularly deciphering the sign of crossings – and labourious. Alternately, gel electrophoresis will stratify nicked DNA knots/catenanes of a given molecular mass and charge. Typically, the distance a given knot or catenane migrates through the gel is proportional to the minimal crossing number (MCN, defined below), with knots of greater MCN migrating more rapidly than those with lesser MCN⁷⁻⁹ ^b But there are 1,701,936 knots with $MCN \leq 16$, so a better stratification is needed to positively identify a particular knot.¹⁰ Recent work has shown that 2-dimensional gel electrophoresis can separate some prime knots with the same MCN.⁵ However, there is no clear relationship that determines relative migration of knots with the same MCN in the second dimension.

For DNA of a given length, (1-dimensional) gel electrophoresis can separate some knots with the same MCN. For example, the 5 and 7-crossing torus knots migrate more slowly than the corresponding 5 and 7-crossing twist knots.^{7,11} This has not generalized, although recent experiments indicate that knots/catenanes may migrate linearly with respect to the average

^a Email address of the corresponding author: d.buck@imperial.ac.uk

^b However there are gel conditions where, for example, the unknot will migrate ahead of the trefoil.

crossing number (ACN) of a particular conformation – the *ideal configuration*^c of the knot or catenane.^{11,13,14}

For gel electrophoresis, one must also construct an appropriate knot ladder as a control to determine the exact DNA knot or catenane, since adjacent bands determine only relative MCN or ACN, not precise values. While this can be done in some cases (*e.g.* T4 topoisomerase will produce a ladder of twist knots¹⁵) generating such a ladder of known knots/catenanes from DNA of the same length and sequence as the unknown knots is highly nontrivial.

Thus topological techniques, such as those presented here, can aid experimentalists in characterizing DNA knots and catenanes.

In this work, we focus on knots and catenanes that arise from site-specific recombination. *Site-specific recombinases* mediate such a rearrangement of the genome (see *e.g.*^{16,17} for a more thorough introduction). Loosely, the recombinases bind and synapse two small segments of DNA, then cleave (by nucleophilic displacement of a DNA hydroxyl by a protein side chain), exchange and reseal the backbones, before releasing the DNA. The result of site-specific recombination can be excision, integration, or inversion of DNA. This corresponds to a wide variety of physiological processes, including integration of viral DNA into the host genome, bacterial gene replication and plasmid copy number control. If the substrate DNA contains supercoils, or if synapsis introduces DNA crossings, these crossings can become knot or catenane nodes in the product.

Topological techniques have already played a significant role in characterizing knotted and catenated products of site-specific recombination. For example, several approaches have been developed to determine a particular DNA knot or catenane type, including utilizing the node number for knots,² the Jones polynomial for catenanes,¹⁸ Schubert’s classification of 4-plats¹⁹ and the HOMFLY polynomial.²⁰ Perhaps most famously, Ernst and Sumners have developed the tangle model of recombination to describe the action of particular site-specific recombinases in terms of tangle sums.²¹ The tangle model has since been used to determine various features of protein-DNA interactions for a number of specific proteins.^{22–31}

With the exception of ²⁹ discussed below, the previous topological treatments began with the precise, biologically determined knot or catenane types of (at least some of) the products. This input was then harnessed in topological arguments that probed various features of the pathway and/or mechanism. Here we consider the alternate paradigm: given a few assumptions about the mechanism, we predict which knots/catenanes are putative products.

More specifically, rather than focusing on a specific recombinase as many earlier studies have done, we present a topological model that predicts which knots/catenanes can occur as products of site-specific recombination *in general*. We do this by describing the topology of how DNA knots and catenanes are formed as a result of a single – or multiple rounds of processive – recombination event(s), given an unknot, unlink, or $(2, m)$ -torus knot or catenane substrate. Our model relies on three assumptions, and we provide biological evidence for each. Given these assumptions, this model predicts that products arising from site-specific recombination must be members of a single family of products (illustrated in Figure 4).^d In ³² we provide the technical proofs for the model developed here, whose nascent form we sketched in.³³

^c *Ideal geometric configurations* of knots or catenanes are the trajectories that allow maximal radial expansion of a virtual tube of uniform diameter centered around the axial trajectory of the knot.^{12,13}

^d Note *all* figures represent (the axis of) duplex DNA.

This paper complements earlier work of Sumners *et al.*,²⁹ which used the tangle model and several biologically reasonable assumptions to solve tangle equations. They then determined which 4-plat knots and catenanes arise as a result of (possibly processive) site-specific recombination on the unknot for the serine subfamily of recombinases. (See below for a discussion of the 2 subfamilies.) For the particular case of the recombinase Gin, they considered the knots $3_1, 4_1, 5_2$ or 6_1 as substrates as well as unknotted substrates. The current work goes further in several ways. In addition to an unknotted substrate for a generic recombinase, we allow substrates that are unlinks with one site on each component, as well as $(2, m)$ -torus knots and catenanes. Also, our assumptions are based exclusively on the biology of the recombination process. In particular, we do not assume the tangle model holds or that the products must be 4-plats. This is particularly important as recombination has been seen to produce knots and catenanes which are connected sums, and thus not 4-plats (see Table 2).

The paper is organized as follows. In Section 2, we state the three assumptions of our model about the recombinase-DNA complex, and give supporting experimental evidence for each. In Section 3, we show that, given an unknot, unlink or $(2, m)$ -torus knot or catenane substrate, all possible knotted or catenated products fall into a single characterized family. We also consider the (common) case of substrates which are $(2, m)$ -torus knots and catenanes whose products have minimal crossing number $m + 1$, and show that the product knot or catenane type is even more tightly prescribed. In Section 4, we first show that for substrates which are unknotted, unlinked, or $(2, m)$ -torus knots or catenanes all known products fall within our family. (The technical proofs of the results in Sections 3 and 4 can be found in.³²) We then apply our results to narrow the possible knot or catenane type for previously uncharacterized experimental data.

2.1. Background and Terminology. We define standard DNA as that which is covalently closed, duplex and plectonemically supercoiled possibly with branch points. Roughly speaking a circular DNA molecule is plectonemically supercoiled if there is a second order helix formed by the DNA axis itself (see¹ for a more complete description). It is believed that supercoiled DNA is the typical form of DNA *in vivo*.¹⁸ Branched DNA structures within supercoiled plasmids *in vitro* have been visualized by electron microscopy.^{18,34} Additionally atomic force microscopy *in situ* illuminates branched plectonemic superhelices at physiological conditions.³⁵ *In vivo*, there is evidence from several more indirect experiments that branched supercoiled DNA is ubiquitous, e.g.³⁶

During site-specific recombination, a recombinase dimer first binds to each of two specific DNA sites of approximately 20-30 basepairs. We refer to these sites as the *crossover sites*. The two crossover sites are then brought together within a *recombinase complex*, B : the smallest convex region containing the four bound recombinase molecules and the two crossover sites. So B is a *topological ball* (*i.e.*, it can be deformed to a round ball). The crossover sites can be located either on the outside, separated by the catalytic domains, (*e.g.* with $\gamma\delta$ and Tn3 resolvase), or inside the 4 recombinase subunits.³⁶⁻⁴³ We will use the term *recombinase-DNA complex* to refer to B together with the substrate. If the recombinase complex meets the substrate in precisely the two crossover sites then we say the recombinase complex is a *productive synapse*.

The existence of a productive synapse for recombinases is in contrast with transposases whose enhancer sequences are intertwined with the active transposition sites, *e.g.*,³ preventing the existence of a productive synapse. Figure 1 demonstrates two examples where the recombinase complex B is a productive synapse, and one where B is not.

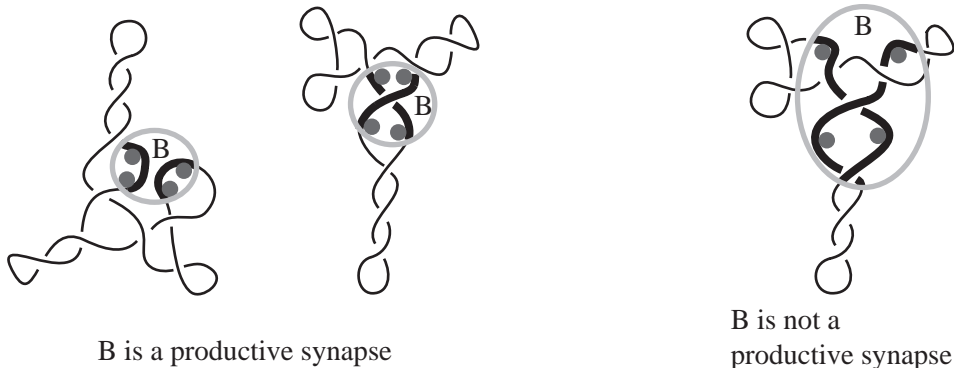


FIGURE 1. Productive Synapse. We require that the recombinase complex is a *productive synapse*, that is the recombinase complex meets the substrate in precisely the two crossover sites. The two examples on the left have a productive synapse and the one on the right does not. The crossover sites are highlighted in black.

Site-specific recombinases fall into two families – the serine (also known as the resolvase) and tyrosine (also known as the integrase) recombinases – based on sequence homology and catalytic residues.¹⁷ The serine and tyrosine recombinases also differ in their mechanism of cutting and rejoining DNA at the crossover sites. Both families are large: a phylogenetic analysis has been performed on 72 serine recombinases⁴⁴ and a recent iterative PSI-BLAST search documents approximately 1000 related sequences of putative tyrosine recombinases.⁴⁵

The large, diverse family of serine recombinases is comprised of resolvases (such as Tn3 and $\gamma\delta$), invertases (such as Gin, Hin, Pin, and Min), large serine recombinases (also called large resolvases) and IS elements.⁴⁴ These recombinases may trap a fixed number of supercoils before initiating recombination.¹⁷ For example, Tn3 resolvase requires three negative supercoils to be trapped by the binding of (non-active) resolvase molecules. These trapped supercoils (outside of the recombinase complex) together with the recombinase complex itself are known as the *synaptic complex*.^{46,47} Likewise, the invertases also require a fixed number of supercoils trapped outside the recombinase complex. Rather than using additional recombinase molecules, they rely on accessory proteins and enhancer sequences, which facilitate the organization of a unique stereospecific synapse that promotes DNA cleavage. (In the Hin and Gin systems, these bound supercoils, together with the recombinase complex, are referred to as the *invertasome*.^{36,48}) With serine recombinases, recombination proceeds through a concerted 4-strand cleaving and rejoining reaction.¹⁷ Serine recombinases can perform multiple rounds of strand exchange before releasing the DNA, in a process known as *processive recombination*.

In contrast, tyrosine recombinases first cleave, exchange and reseal two sugar-phosphate backbones. The DNA-protein complex then proceeds through an intermediary structure (a

Holliday junction) before repeating the process with the other two DNA backbones.^{17,49} Most tyrosine recombinases, including Flp, λ Int and Cre, tolerate varying number of supercoils outside of the recombinase complex. However, there are exceptions, most notably XerCD, which trap a fixed number of supercoils using accessory proteins before initiating cleavage.⁵⁰ Like serine recombinases, tyrosine recombinases can also employ accessory proteins to help assemble the synaptic complex, and to drive the overall reactions (*e.g.* λ Int and XerCD).^{43,50}

Finally, we shall use the following knot theoretic terms. The *components* of a catenane are the separate rings of the catenane. A knot is considered to be a catenane with only one component. A $(2, m)$ -torus knot or catenane is one which can be drawn so that all of its crossings occur as a row of m twists, as illustrated in Figure 2. We will denote a knot or catenane of this form by $T(2, m)$. Any such knot or catenane is the boundary of a twisted annulus. If m is odd then $T(2, m)$ is a knot and if m is even then $T(2, m)$ is a catenane. Finally, given two knots or catenanes K and J , their *connected sum*, written $K \# J$, is obtained by removing a trivial arc from each and gluing the resulting two endpoints of K to the two endpoints of J without introducing any additional knotting. Figure 4, subfamilies 4 and 5 give examples of connected sums.

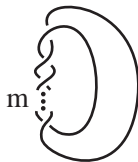


FIGURE 2. Torus Knots and Catenanes. A $(2, m)$ -torus knot or catenane has this form. The model considers substrates that are unknots, unlinks or $(2, m)$ -torus knot or catenanes.

3. THE ASSUMPTIONS OF OUR MODEL

We begin with a fixed recombinase and an unknot, unlink, or $T(2, m)$ substrate. If the substrate is an unlink then we assume that one site is on each component, as otherwise this case reverts back to a single unknotted substrate. We make three biological assumptions about the recombinase-DNA complex, stated in both biological and mathematical terms, and provide experimental evidence for each.

Let J denote the substrate(s) after synapsis, and recall that B denotes the recombinase complex.

Assumption 1: The recombinase complex is a productive synapse, and there is a projection of the crossover sites which has at most one crossing between the sites and no crossings within a single site.

Evidence for Assumption 1:

We present a variety of experimental data that suggests that the recombinase complex is sufficiently dense both to form a proper productive synapse and to preclude extraneous crossings.

Most convincingly, recent crystal structures of several recombinase complexes support both of these assumptions.^{40–43,51} Additionally, structures of a single site synapsed with either a

dimer or monomer indicate that there are no crossings within an individual site.^{52–54} Furthermore, structures of two intermediate complexes – a synaptic $\gamma\delta$ resolvase tetramer covalently linked to two *cleaved* DNAs, and the Flp recombinase-Holliday junction complex have at most one crossing between sites and none within a site.^{39,42} Thus the large-scale conformational changes necessary to unwrap crossings during the reaction imply it is unlikely that the crossover sites contain additional crossings at synapsis or that a productive synapse does not exist.

Also, there are significant DNA-protein and protein-protein interactions that appear to prevent additional crossings and extraneous strands from piercing the recombinase complex. With tyrosine recombinases, each domain flanking the crossover site DNA inserts a helix into a major groove, and the highly conserved C-terminal domain interacts with consecutive minor and major grooves on the opposite face of DNA.¹⁷ With serine recombinases, DNA binding involves the conserved H-T-H domain, and a DNA binding domain on the C or N terminus of the protein.^{39,44,54} For both families, there are also significant protein-protein interfaces with the other proteins in an assembled tetrameric complex. Additionally, DNA itself has a geometric diameter of 2nm, and, depending on the ionic conditions, a much greater electrostatic diameter (*e.g.* \simeq 5nm at physiological conditions).⁵⁵

Additional biochemical experiments support both the existence of a productive synapse and a bound on the crossings between or within the sites. Atomic force microscopy of both the Cre and Flp productive synapses concur with the conclusions drawn from the crystal structures of the resolvases and integrases.²⁶ Also the architecture of the $\gamma\delta$ resolvase recombinase complex has been determined to be a productive synapse with a single crossing, in experiments using constrained DNA.³⁷ Furthermore, solution structures from neutron and Xray scattering data of hyperactive Tn3 resolvase mutants show that a productive synapse exists, and that there is a projection of the sites with at most one crossing between sites and no crossings within a single site.³⁸ Additionally, recent cyclization experiments indicate that dimers of Flp and Cre each bend the DNA sites upon binding, but not enough to introduce a crossing within a single site.⁵⁶ The steric and electrostatic constraints mentioned above imposed by the short length of the sites also putatively limit crossings between and/or within the sites.

Finally, we note that for recombinases that utilize accessory proteins, we recall that an accessory site or enhancer sequence is neither a crossover site nor a part of a crossover site. Thus in order for our assumption to hold, if a recombinase requires an enhancer sequence, then it must be sequestered from the crossover sites. In particular, we claim that when the enhancer or an accessory site loops around to form a specific recombinase complex all crossings are trapped outside of the complex, even though the recombinases might interact directly with the enhancer sequence. Supporting evidence is twofold. Firstly, the recent λ Int-DNA complex crystal structure includes the accessory sites, and it is clear that a productive synapse exists and has the required limited number of crossings.⁴³ Additionally, support for the invertase family comes both from detailed biochemical experiments of the Hin system. The standard molecular model of the Hin invertasome, based on the cross-linked structure of the Hin-DNA co-complex, has two Fis dimers bound to the enhancer sequence and two Hin dimers bound at the recombination sites.³⁶ According to this model, the enhancer sequence is sequestered from the crossover sites, and the crossover sites are not interwound.

All of the above evidence indicates that it is biologically reasonable to assume that a given recombinase-DNA complex satisfies Assumption 1.

Assumption 2: The productive synapse does not pierce through a supercoil or a branch point in a nontrivial way. Also, no persistent knots are trapped in the branches of the DNA on the outside of the productive synapse.

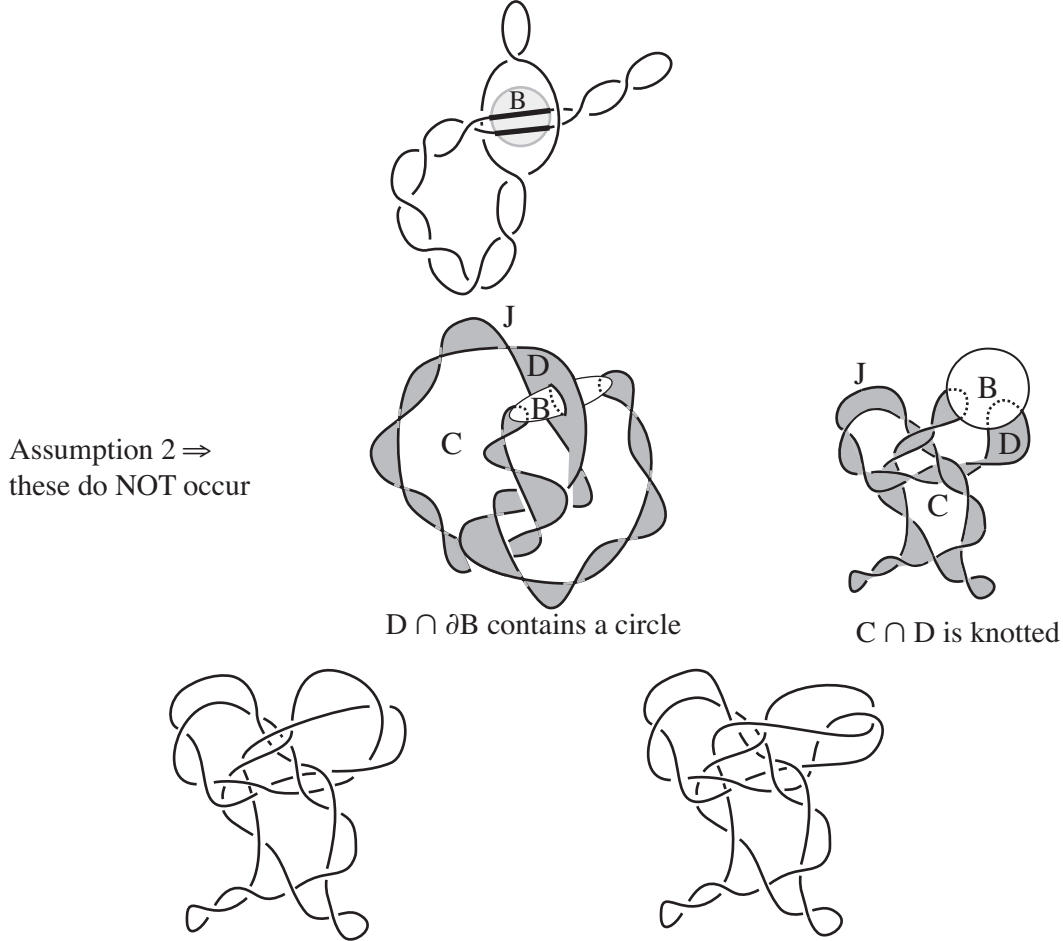


FIGURE 3. Scenarios for Assumption 2. *Above:* the productive synapse B trivially pierces through a supercoil. This is allowed. *Middle Left:* B pierces through a supercoil in a nontrivial way. *Middle Right:* A knot is trapped in the branches on the outside of B . These are forbidden. *Below:* If knots were trapped within the branches of the substrate(s) on the outside of the productive synapse, then recombination would at least occasionally produce *satellite* knots and catenanes such as these. These are forbidden.

Evidence for Assumption 2:

A variety of microscopy studies support Assumption 2. Atomic force microscopy revealed that at physiological conditions, supercoiled DNA adopts a compact plectonemic configuration with close juxtaposition of DNA segments in the loops, which makes it unlikely that a supercoiled domain could be penetrated.^{35,57} Under conditions that minimize intersegmental repulsion, electron and scanning force microscopy studies both demonstrate that opposing

segments of interwound supercoiled DNA are frequently close together.^{18,34,58} Probabilistically then, it is thus unlikely that either a supercoiled domain or B could pierce through a supercoil.

Also, experimental work coupled with MMC simulations indicate that on average supercoiled DNA helices are separated by 10nm – only 5 times the width of the DNA diameter itself.^{18,55} Since, as discussed above, the *electrostatic* diameter of DNA in physiological conditions is closer to 5nm, it seems quite unlikely that the productive synapse could pass through a supercoil (as in Figure 3).^{59–61} Brownian dynamics simulations of site juxtaposition support these findings.⁶²

Additionally, the probability of one duplex (linear) invading a supercoiled domain has been shown to be quite low, by both experiments and MMC simulations.⁵⁹ This frequency may be even lower if a supercoiled domain, rather than linear duplex, invades another supercoiled domain. Therefore trapping a persistent knot within the branches of the DNA during synapsis is also unlikely.

Also the steric and electrostatic constraints arising from protein-DNA interactions discussed in the evidence for Assumption 1 would appear to preclude piercing of the productive synapse by non-site DNA.

Finally, if persistent knots could be trapped within the branches of the substrate(s) on the outside of the productive synapse then we would expect to see (at least occasionally) “doubly knotted” products like those illustrated in Figure 3 (such knots and catenanes are known as *satellites*). However, no products like these have thus far been observed (see Table 2). This indicates that knotting of the branches is unlikely to occur.

All of the above evidence indicates that it is biologically reasonable to assume that a given recombinase-DNA complex satisfies Assumption 2.

Next, we state Assumption 3, which addresses the mechanism of recombination for serine and tyrosine separately.

Assumption 3 for Serine Recombinases: Serine recombinase performs recombination via the “subunit exchange mechanism.” This mechanism involves making two simultaneous (double-stranded) breaks in the sites, rotating opposites sites together by 180° within the productive synapse and resealing opposite partners. In processive recombination, each recombination event is identical.

Assumption 3 for Tyrosine Recombinases: After recombination mediated by a tyrosine recombinase, there is a projection of the crossover sites which has at most one crossing.

Evidence for Assumption 3:

Serine Recombinases: A large number of *in vitro* topology studies performed on DNA invertases and resolvases have provided solid support for the “subunit exchange mechanism,” where one set of recombinase subunits, each covalently associated with the 5’ ends of the cleaved recombination sites, switches places, resulting in a 180° rotation of DNA strands. (See ¹⁷ and references therein.) This is supported by a recent crystal structure of a synaptic tetramer of $\gamma\delta$ resolvase covalently catenated to two cleaved DNA molecules, indicating a subunit rotation of 180° .³⁹

Additional experiments involving Tn3, Hin and Gin lends credence to the idea that each round of processive recombination acts identically.^{2,48,63} For example, Heichman *et al* demonstrate that there are multiple rounds of exclusively clockwise subunit rotation of one set of Hin subunits after DNA cleavage.⁴⁸

Tyrosine Recombinases: While there are no known post-recombinant crystal structures, there are synaptic intermediary crystal co-complexes for Flp,⁴² Cre,⁴¹ and λ Int⁴³ (this also including accessory sites in addition to the typical λ Int crossover sites). These structures indicate that at the earlier stages of recombination – namely after the first cleavage, exchange or within a Holliday junction intermediate – there exists a projection with at most one crossing. They also highlight particular features of the productive synapse that may impede the large-scale conformational changes needed to introduce crossings.

As mentioned above, the protein-DNA interface is a large hydrogen-bonded network. Flp, Cre and λ Int all form a C-shaped clamp around the DNA substrate, and the C-terminal domains interact with consecutive minor and major grooves on the opposite face of the DNA.¹⁷ Additionally, there are significant protein-protein interactions, *e.g.* the catalytic domains interact by swapping part of the C-terminus with a neighbouring protomer.

Also, the post-recombinant complex is formed from the Holliday junction intermediate by, first an isomerization of the intermediary complex so that the inactive monomers become active and *vice versa*, and then a repeated strand cleavage where the new 5' ends migrate over and attack their partners' 3' phosphotyrosine linkages. This second round of strand transfer completes the reaction. Particularly given the two-fold symmetry of the reaction, it thus seems unlikely that in the final stage of recombination there is enough motion of the DNA arms to generate multiple additional crossings between sites or a crossing within a single site.

In vitro studies also suggest that tyrosine recombinases that mobilize the gene cassettes of integrons may preferentially bind DNA hairpins, which would constrain the number of crossings.⁶⁴ Finally, given the steric and electrostatic constraints of short DNA arms discussed for Assumption 1, it is probable that there exists a projection of the sites containing at most one crossing between sites and no crossings of a single site within the post-recombinant complex.

All of the above evidence indicates that it is biologically reasonable to assume that a given recombinase-DNA complex satisfies Assumption 3.

4. RESULTS

4.1. All products of unknots, unlinks or $(2, m)$ torus links substrates fall within a single family. In this section, we suppose that the substrate is an unknot, an unlink, or $T(2, m)$ and that all three of our assumptions hold for a particular recombinase-DNA complex. Below we state Theorems 1 and 2, whose technical proofs can be found in.³² These theorems demonstrate that all knotted and catenated products brought about by that recombinase are in the family of knots and catenanes illustrated in Figure 4.

Observe that p , q , r , and s can be positive, negative, or zero. Furthermore, by letting p , q , r , and/or s be 0 or 1 in Figure 4 as appropriate, we obtain the five subfamilies illustrated in Figure 4. Observe that if $q = 0$, $r = 1$, and $s = -1$, then we have a $T(2, p)$ together with an uncatenated trivial component. This possibility occurs as a member of Subfamily 4. Thus the knots and catenanes in these subfamilies are all possible products of recombination as

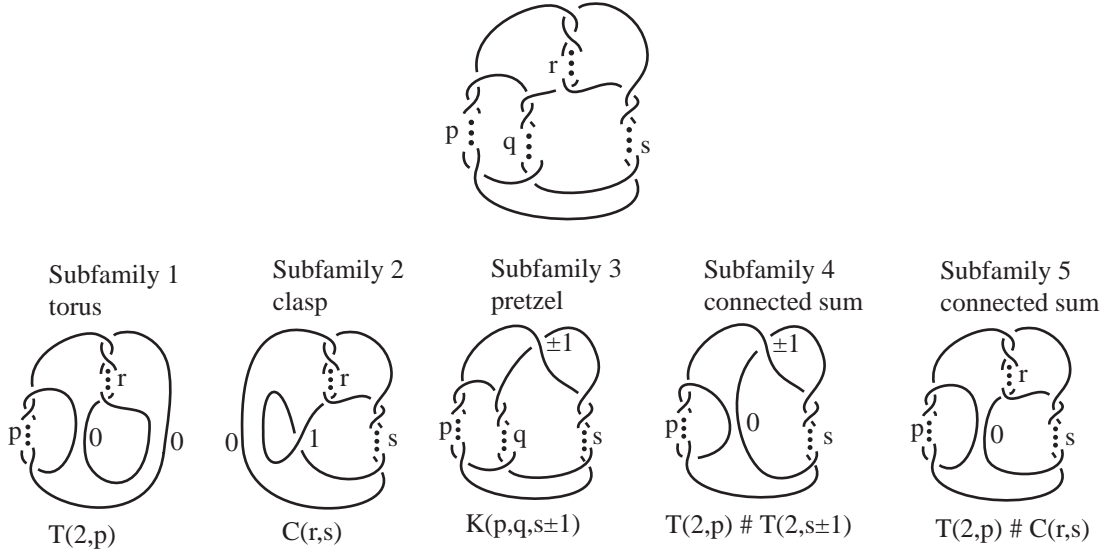


FIGURE 4. Product Family. Given an unknotted, unlinked or torus knot or catenane substrate, then all products are of this type. The five subfamilies are indicated below.

specified in Theorems 1 and 2. We use the notation $C(r, s)$ for a knot or catenane consisting of one row of r crossings and a non-adjacent row of s crossings (illustrated in Subfamily 2). Note that if r or s equals 2, then $C(r, s)$ is in the well known family of *twist* knots and catenanes. We use the notation $K(p, q, r)$ for a pretzel knot or catenane with three non-adjacent rows containing p crossings, q crossings, and r crossings (illustrated in Subfamily 3, where $r = s \pm 1$). Note, by *non-adjacent* rows of r and s crossings we mean that the two rows cannot be considered as a single row of $r + s$ crossings.

Theorem 1. *Suppose that Assumptions 1, 2, and 3 hold for a particular tyrosine recombinase-DNA complex. If the substrate is an unknot then the only non-trivial products are $T(2, n)$ and $C(2, n)$. If the substrate is an unlink, then the only non-trivial product is a Hopf link. If the substrate is $T(2, m)$ then all of the non-trivial products are contained in the family illustrated in Figure 4.*

Theorem 2. *Suppose that Assumptions 1, 2, and 3 hold for a particular serine recombinase-DNA complex. If the substrate is an unknot then the only non-trivial products are $T(2, n)$ and $C(p, q)$. If the substrate is an unlink, then the only non-trivial product is $T(2, n)$. If the substrate is $T(2, m)$ then all non-trivial products are in the general family illustrated in Figure 4.*

Table 1 summarizes the non-trivial products predicted by Theorems 1 and 2. Recall that $C(r, s)$ is a twist knot or catenane if r or s equals 2.

4.2. Products are more tightly restricted when recombination adds 1 crossing. Knots and catenanes have been tabulated according to the fewest number of crossings with

Recombinase Type	Substrate Topology	Non-trivial Products
Tyrosine	unknot	$T(2, n), C(2, n)$
	unlink	Hopf link = $T(2, 2)$
	$T(2, m)$	Family of Figure 4
Serine	unknot	$T(2, n), C(p, q)$
	unlink	$T(2, n)$
	$T(2, m)$	Family of Figure 4

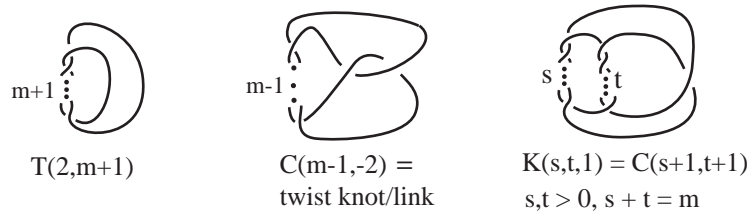
TABLE 1. Non-trivial products predicted by our model.

which they can be drawn (see the tables ⁶⁵ and ¹⁰). This number of crossings is called the *minimal crossing number* of the knot or catenane, and is denoted by MCN. For example a $T(2, 2)$ (also known as a Hopf link) has MCN = 2 and $T(2, 3)$ (also known as a trefoil knot) has MCN = 3. In fact, $\text{MCN}(T(2, m)) = m$ for any positive integer m . Gel electrophoresis can be used to determine the MCN of a product (see *e.g.*⁹).

It is often the case that recombination adds a single crossing to the MCN of a knotted or catenated substrate, *e.g.*⁶⁶ If the substrate is $T(2, m)$ and the product has $\text{MCN} = m + 1$, then we can further refine the results of Theorems 1 and 2 to determine more specific possibilities for the products, the technical details of which can be found in.³² The conclusion of the theorem is illustrated in Figure 5.

Theorem 3. *Suppose that Assumptions 1, 2, and 3 hold for a particular recombinase-DNA complex with substrate $T(2, m)$, with $m > 0$. Let L be the product of a single recombination event, and suppose that $\text{MCN}(L) = m + 1$. Then L is either $T(2, m + 1)$, $C(m - 1, -2)$, or $K(p, q, 1)$ with $p, q > 0$ and $p + q = m$.*

Furthermore, $K(p, q, 1)$ is a knot if and only if at least one of p and q is odd.


 FIGURE 5. Restricted Products. These are the only possible products, if the substrate is $T(2, m)$ and the product has $\text{MCN} = m + 1$.

5. APPLICATIONS

5.1. All Characterized Recombinant Products are in the Predicted Family. Table 3 summarizes the known products of recombinases starting with substrates which are unknots, unlinks, or $T(2, m)$. As shown in Table 3, all products listed have a projection in the form of Figure 4. This provides further confirmation of the validity of our model.

Note that this table does not describe every product of site-specific recombination – *e.g.* Tn3 acting on the twist knot 4_1 yields the product 5_1^2 ,⁴⁶ and mutant Hin acting on a

Recombinase	Substrate	Products	Subfamily	Reference
Cre	0_1 (Inverted)	$0_1, 3_1, 5_1, 7_1, 9_1, 11_1$	1	⁶⁷
	-2_1^2 (Direct)	$0_1, 3_1$	1	²⁴
	0_1 (Direct)	$0_1^2, 2_1^2, 4_1^2, 6_1^2, 8_1^2, 10_1^2, 12_1^2$	1	⁶⁷
Flp	0_1 (Inverted)	$0_1, 3_1, 5_1, 7_1, 9_1, 11_1$	1	^{24, 68}
	0_1 (Direct)	$0_1^2, 2_1^2, 4_1^2, 6_1^2, 8_1^2, 10_1^2, 12_1^2$	1	^{24, 68}
λ Int	0_1 (PB inverted)	$0_1, 3_1, 5_1, 7_1, 9_1, 11_1, 13_1, 15_1, 17_1, 19_1$	1	⁶⁹
	0_1 (PB direct)	$4_1^2, 6_1^2$	1	⁶⁹
	-2_1^2 (PB direct)	$5_2, 7_2, 9_2, 11_2$	$2 = C(2, s)$	²⁴
	0_1 (LR inverted)	$3_1, 5_1, 7_1$	1	²⁴
	0_1 (LR direct)	4_1^2	1	²⁴
Xer	-2_1^2 (LR direct)	$0_1, 3_1, 5_2, 7_2, 9_2, 11_2, 13_2$	$1, 2 = C(2, s)$	²⁴
	0_1 (Direct)	4_1^2	1	⁵⁰
Gin (inverted)	0_1	0_1	1	⁶³
	0_1	3_1	1	⁶³
	3_1	4_1	1	⁶³
	3_1	$3_1 \# 3_1$	4	⁶³
	4_1	5_2	1	⁶³
Gin (direct)	0_1	3_1	1	⁶³
	3_1	5_2	2	⁶³
	3_1	$3_1 \# 3_1$	4	⁶³
	5_2	7_2	2	⁶³
Gin (mutant, inv'd)	0_1	$3_1, 4_1, 5_1, 5_2, 7_1, 7_2, 9_1, 9_2$	1, 2	⁷⁰
Hin (inverted)	0_1	0_1	1	³⁶
	0_1	3_1	1	³⁶
	3_1	4_1	2	³⁶
	4_1	5_2	2	³⁶
Hin (mutant)	0_1	$3_1, 5_2, 3_1 \# 3_1$	1, 2, 4	⁴⁸
Tn3, $\gamma\delta$ (direct)	0_1	2_1^2	1	⁴⁶
	2_1^2	4_1	2	⁴⁶
	4_1	5_1^2	2	⁴⁶
	5_1^2	$4_1 \# 2_1$	4	⁴
	5_1^2	6_2	2	⁴⁶

TABLE 2. As predicted by our model, all characterized products of site-specific recombination on supercoiled unknotted, unlinked or $(2, n)$ -torus knot or catenane substrates fall within our single family (see Figure 4).

heterogenous population of 3_1 and twist knots 5_2 yields (double and, in multiple rounds of recombination, triple) connected sums: $3_1 \# 3_1, 3_1 \# 5_2, 5_2 \# 5_2$ and $3_1 \# 3_1 \# 5_2$ ³⁶ – but we do not consider these (twist knot or two separate knots) as substrates in our model.

5.2. Applications to Uncharacterized Recombinant Products. We now turn our attention to several recombination systems whose products are unclassified beyond minimal crossing number. We use our model, together with results about minimal crossing number, to prove that the product knot or catenane type is tightly prescribed, and apply this new result to the previously uncharacterized experimental data.

For each, we discuss how our model can help to restrict the knot types of these products.

Xer: Using a plasmid with both λ Int and Xer sites, Bath *et al* generated the catenanes 6_1^2 and 8_1^2 as products of λ recombination.⁶⁶ These were then used as the substrates for Xer recombination, yielding a knot with MCN=7 and a knot with MCN=9, respectively. These products have not been characterized beyond their minimal crossing number. There are seven knots with MCN=7 and 49 knots with MCN=9.

Theorem 3 significantly reduces the number of possibilities for each of these products. In particular, it follows from Theorem 3 that the 7-crossing products of Xer must be $7_1 = T(2, 7)$, $7_2 = C(5, -2)$ or $7_4 = K(3, 3, 1)$; and the 9-crossing products of Xer must be $9_1 = T(2, 9)$, $9_2 = C(7, -2)$, or $9_5 = K(5, 3, 1)$. Observe that all of these knots are 4-plats. This demonstrates how our model complements earlier work of,²⁵ which assumes all products must be 4-plats and hence only considers 7-crossing products (since only half of the 9-crossing knots are 4-plats). In ⁷¹ we use our model together with tangle calculus to completely classify all tangle solutions to these λ Int-Xer equations.

Cre111: Abremski and collaborators created the mutant Cre111, which yields products topologically distinct from those of wild-type Cre.⁷² When Cre111 recombines a supercoiled substrate, the knotted and catenated products are, in their conditions, significantly more complex than those produced by wild-type Cre.

These knots and catenanes have thus far been uncharacterized. However, our Theorem 1 predicts that these knots and catenanes must be of the form $T(2, n)$ or $C(2, n)$. Thus, by running these products adjacent to a ladder of torus knots of the same length, one could determine the exact knot or catenane type.

Tn3: Benjamin *et al* constructed a plasmid substrate for Tn3 resolvase with four directly repeated crossover sites.⁷³ After the first round of recombination, electron microscopy reveals the Hopf link $T(2, 2)$ as the primary product. After recombination, products were determined (via high resolution gel electrophoresis (of 7-8 days) followed by electron microscopy, as $T(2, 2) \# T(2, 2)$) and two distinct 4-component catenates. Our model predicts that, with this 4-sited substrate, recombination must proceed from the unknot to the Hopf link $T(2, 2)$. It then utilizes this $T(2, 2)$ catenane as a substrate to yield the product the connected sum of two Hopf links $T(2, 2) \# T(2, 2)$ (see Table 4.1). This connect sum is then the substrate for the products of 4-component catenanes, but is not one of the substrates that we consider. This would be akin to Tn3 performing multiple rounds of *distributive* recombination on a substrate with only 2 crossover sites. The current work thus supports Benjamin *et al*'s hypothesis of neighbouring-site recombination.

6. CONCLUDING REMARKS

In this paper we have developed a model of how DNA knots and catenanes are produced as a result of a recombinase acting on an unknot, unlink, or $(2, n)$ -torus knot or catenane substrate. Our model is based on three explicitly stated biological assumptions about site-specific recombination, and we have provided biological evidence for each. It follows from our model that all knotted or catenated products of such enzyme actions will be in the family of Figure 4, as described in Theorems 1 and 2.

As mentioned above, the minimal crossing number (MCN) of a DNA knot or catenane can be determined experimentally.⁹ For small values of the MCN there are not many knots or catenanes with a given value. However, the number of knots and catenanes with MCN = n

grows exponentially as a function of n ,⁷⁴ and there are 1,701,936 knots with $\text{MCN} \leq 16$.¹⁰ So knowing the MCN is not sufficient to determine the knot or catenane.

However the total number of knots and catenanes in the family of Figure 4 grows linearly with n^3 .³² So the proportion of all knots and catenanes which are contained in our family decreases exponentially as n increases. Thus, knowing the MCN of a product and knowing that the product is in one of our families allows us to significantly narrow the possibilities for its knot or catenane type. The model described herein thus provides an important step in characterizing DNA knots and catenanes which arise as products of site-specific recombination.

7. ACKNOWLEDGEMENTS

The authors wish to thank J. Arsuaga, I. Grainge, R. Harshey, M. Jayaram, A. Stasiak, D.W. Summers, A. Vologodskii and S. Whittington for helpful conversations. DB was partially supported by NSF Grant # DMS-0102057. EF was partially supported by an Association for Women in Mathematics Michler Collaborative Research Grant.

REFERENCES

- [1] A. Maxwell and A. D. Bates. *DNA Topology*. Oxford University Press, 2005.
- [2] N. R. Cozzarelli, M. A. Krasnow, S. P. Gerrard, and J. H. White. A topological treatment of recombination and topoisomerases. *Cold Spring Harb Symp Quant Biol*, 49:383–400, 1984.
- [3] S. Pathania, M. Jayaram, and R. M. Harshey. Path of DNA within the Mu transpososome. transposase interactions bridging two Mu ends and the enhancer trap five DNA supercoils. *Cell*, 109(4):425–36, 2002.
- [4] M. A. Krasnow, A. Stasiak, S. J. Spengler, F. Dean, T. Koller, and N. R. Cozzarelli. Determination of the absolute handedness of knots and catenanes of DNA. *Nature*, 304(5926):559–60, 1983.
- [5] S. Trigueros, J. Arsuaga, M. E. Vazquez, D. W. Summers, and J. Roca. Novel display of knotted DNA molecules by two-dimensional gel electrophoresis. *Nucleic Acids Res*, 29(13):E67–7, 2001.
- [6] E.L. Zechiedrich and N.J. N.J. Crisona. *Methods in Molecular Biology: DNA Topoisomerase Protocols*, volume 1 of *M. Bjornsti and N. Osheroff (eds)*, chapter Coating DNA with RecA protein to distinguish DNA path by electron microscopy, pages 99–108. Humana Press, Totowa, NJ, 1989.
- [7] V. Katritch, J. Bednar, D. Michoud, J. Dubochet, and A. Stasiak. Electrophoretic mobility of DNA knots. *Nature*, 384(6605):122, 1996.
- [8] O. Sundin and A. Varshavsky. Terminal stages of SV40 DNA replication proceed via multiply intertwined catenated dimers. *Cell*, 21(1):103–14, 1980.
- [9] S. D. Levene and H. Tsen. Analysis of DNA knots and catenanes by agarose-gel electrophoresis. *Methods Mol Biol*, 94:75–85, 1999.
- [10] Jim Hoste, Morwen Thistlethwaite, and Jeff Weeks. The first 1,701,936 knots. *Math. Intelligencer*, 20(4):33–48, 1998.
- [11] A. V. Vologodskii, N. J. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet, and A. Stasiak. Sedimentation and electrophoretic migration of DNA knots and catenanes. *J Mol Biol*, 278(1):1–3, 1998.
- [12] A. Y. Grosberg, A. Feigel, and Y. Rabin. Flory-type theory of a knotted ring polymer. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics*, 54(6):6618–6622, 1996.
- [13] V. Katritch, J. Bednar, D. Michoud, R. G. Scharein, J. Dubochet, and A. Stasiak. Geometry and physics of knots. *Nature*, 384:142–145, 1996.
- [14] B. Laurie, V. Katritch, J. Dubochet, and A. Stasiak. Geometry and physics of links. *Biophysical Journal*, 74, 1998.
- [15] S. A. Wasserman and N. R. Cozzarelli. Supercoiled DNA-directed knotting by T4 topoisomerase. *J Biol Chem*, 266(30):20567–73, 1991.
- [16] N. Craig, R. Craigie, M. Gellert, and A. (eds) Lambowitz. *Mobile DNA*. ASM Press, Washington, DC, 2002.

- [17] N. D. Grindley, K. L. Whiteson, and P. A. Rice. Mechanisms of site-specific recombination. *Annu Rev Biochem*, 75:567–605, 2006.
- [18] T. C. Boles, J. H. White, and N. R. Cozzarelli. Structure of plectonemically supercoiled DNA. *J Mol Biol*, 213(4):931–51, 1990.
- [19] J. H. White and N. R. Cozzarelli. A simple topological method for describing stereoisomers of DNA catenanes and knots. *Proc Natl Acad Sci U S A*, 81(11):3322–6, 1984.
- [20] J. H. White, K. C. Millett, and N. R. Cozzarelli. Description of the topological entanglement of DNA catenanes and knots by a powerful method involving strand passage and recombination. *J Mol Biol*, 197(3):585–603, 1987.
- [21] C. Ernst and D. W. Sumners. A calculus for rational tangles: applications to DNA recombination. *Math. Proc. Cambridge Philos. Soc.*, 108(3):489–515, 1990.
- [22] D. Buck and C. Verjovsky Marcotte. Tangle solutions for a family of DNA-rearranging proteins. *Math. Proc. Cambridge Philos. Soc.*, 139(1):59–80, 2005.
- [23] D. Buck and C. Verjovsky Marcotte. Classification of tangle solutions for integrases, a protein family that changes DNA topology. *Journal of Knot Theory and its Ramifications*, To Appear, arXiv:math/0501173v1.
- [24] N.J. Crisona, R.L. Weinberg, B.J. Peter, D.W. Sumners, and N.R. Cozzarelli. The topological mechanism of phage lambda integrase. *J Mol Biol*, 289(4):747–775, 1999.
- [25] Isabel K. Darcy. Biological distances on DNA knots and links: applications to XER recombination. *J. Knot Theory Ramifications*, 10(2):269–294, 2001. Knots in Hellas '98, Vol. 2 (Delphi).
- [26] A.A. Vetcher, A.Y. Lushnikov, J. Navarra-Madsen, R.G. Scharein, Y.L. Lyubchenko, I.K. Darcy, and S.D. Levene. DNA topology and geometry in Flp and Cre recombination. *J Mol Biol*, 357:1089–1104, 2006.
- [27] I. Grainge, D. Buck, and M. Jayaram. Geometry of site alignment during Int family recombination: Antiparallel synapsis by the Flp recombinase. *J Mol Bio*, 298:749–764, 2000.
- [28] C. Ernst and D. W. Sumners. Solving tangle equations arising in a DNA recombination model. *Math. Proc. Cambridge Philos. Soc.*, 126(1):23–36, 1999.
- [29] D.W. Sumners, C. Ernst, S.J. Spengler, and N. R. Cozzarelli. Analysis of the mechanism of DNA recombination using tangles. *Quarterly Review of Biophysics*, 28(3):253–313, 1995.
- [30] M. Vazquez and D. W. Sumners. Tangle analysis of Gin site-specific recombination. *Math. Proc. Cambridge Philos. Soc.*, 136(3):565–582, 2004.
- [31] M. Vazquez, S. D. Colloms, and D. Sumners. Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three-dimensional topological pathway. *J Mol Biol*, 346(2):493–504, 2005.
- [32] D. Buck and E. Flapan. A topological characterization of knots and links arising from site-specific recombination. *arXiv:0707.3896v1*, 2007.
- [33] D. Buck and E. Flapan. A model of DNA knotting and linking. *OCAMI Studies*, 1(1):75–83, 2007.
- [34] M. Adrian, B. ten Heggeler-Bordier, W. Wahli, A. Z. Stasiak, A. Stasiak, and J. Dubochet. Direct visualization of supercoiled DNA molecules in solution. *EMBO J*, 9(13):4551–4, 1990.
- [35] Y. L. Lyubchenko and L. S. Shlyakhtenko. Visualization of supercoiled DNA with atomic force microscopy in situ. *Proc Natl Acad Sci U S A*, 94(2):496–501, 1997.
- [36] S.K. Merickel and R.C. Johnson. Topological analysis of Hin-catalysed DNA recombination *in vivo* and *in vitro*. *Mol Microbiology*, 51(4):1143–1154, 2004.
- [37] A.E. Leschziner and N.D.F. Grindley. The architecture of the $\gamma\delta$ resolvase crossover site synaptic complex revealed by using constrained DNA substrates. *Mol Cell*, 12:775–781, 2003.
- [38] M. Nollmann, J. He, O. Byron, and W.M. Stark. Solution structure of the Tn3 resolvase-crossover site synaptic complex. *Mol Cell*, 16(1):127–137, 2004.
- [39] W. Li, S. Kamtekar, Y. Xiong, G. Sarkis, N.D.F. Grindley, and T.A. Steitz. Structure of a synaptic resolvase tetramer covalently linked to two cleaved DNAs. *Science*, 309:1210–1215, 2005.
- [40] E. Ennifar, J.E.W. Meyer, F. Buchholz, A.F. Stewart, and D. Suck. Crystal structure of a wild-type Cre recombinase-loxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Res*, 31(18):5449–5460, 2003.
- [41] F. Guo, D.N. Gopaul, and G.D. van Duyne. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, 389(40–46), 1997.

- [42] J. Chen, P. T. Englund, and N. R. Cozzarelli. Changes in network topology during the replication of kinetoplast DNA. *EMBO J*, 14(24):6339–47, 1995.
- [43] T. Biswas, H. Aihara, M. Radman-Livaja, D. Filman, A. Landy, and T. Ellenberger. A structural basis for allosteric control of DNA recombination by λ integrase. *Nature*, 435:1059–1066, 2005.
- [44] M. C. Smith and H. M. Thorpe. Diversity in the serine recombinases. *Mol Microbiol*, 44(2):299–307, 2002.
- [45] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Res*, 25:3389–3402, 1997.
- [46] S. A. Wasserman, J. M. Dungan, and N. R. Cozzarelli. Discovery of a predicted DNA knot substantiates a model for site-specific recombination. *Science*, 229(4709):171–4, 1985.
- [47] W.M. Stark, M.R. Boocock, and D.J. Sherratt. Catalysis by site-specific recombinases. *Trends Genet*, 8:432–439, 1992.
- [48] K.A. Heichman, I.P. Moskowitz, and R.C. Johnson. Configuration of DNA strands and mechanism of strand exchange in the *Hin* invertasome as revealed by analysis of recombinant knots. *Genes Dev*, 5(9):1622–1634, 1991.
- [49] I. Grainge and M. Jayaram. The integrase family of recombinase: organization and function of the active site. *Mol Microbiol*, 33(3):449–56, 1999.
- [50] S. D. Colloms, J. Bath, and D. J. Sherratt. Topological selectivity in Xer site-specific recombination. *Cell*, 88(6):855–64, 1997.
- [51] P. A. Rice and T. A. Steitz. Model for a DNA-mediated synaptic complex suggested by crystal packing of $\gamma\delta$ resolvase subunits. *EMBO J*, 13(7):1514–24, 1994.
- [52] D. G. Blake, M. R. Boocock, D. J. Sherratt, and W. M. Stark. Cooperative binding of Tn3 resolvase monomers to a functionally asymmetric binding site. *Curr Biol*, 5(9):1036–46, 1995.
- [53] D. MacDonald, G. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul. Structural basis for broad DNA-specificity in integron recombination. *Nature*, 440(7088):1157–62, 2006.
- [54] W. Yang and T.A. Steitz. Crystal structure of the site-specific recombinase $\gamma\delta$ resolvase complexed with a 34 bp cleavage site. *Cell*, 82:193–207, 1995.
- [55] A. V. Vologodskii and N. R. Cozzarelli. Conformational and thermodynamic properties of supercoiled DNA. *Annu Rev Biophys Biomol Struct*, 23:609–43, 1994.
- [56] Q. Du, A. Livshits, A. Kwiatek, M. Jayaram, and A. Vologodskii. Protein-induced local DNA bends regulate global topology of recombination products. *J Mol Biol*, 368(1):170–82, 2007.
- [57] L.S. Shlyakhtenko, P. Hsieh, M. Grigoriev, V.N. Potaman, R.R. Sinden, and Y.L. Lyubchenko. A cruciform structural transition provides a molecular switch for chromosome structure and dynamics. *J Mol Bio*, 296(5):1169, 2000.
- [58] D. I. Cherny and T. M. Jovin. Electron and scanning force microscopy studies of alterations in supercoiled DNA tertiary structure. *J Mol Biol*, 313(2):295–307, 2001.
- [59] V. V. Rybenkov, A. V. Vologodskii, and N. R. Cozzarelli. The effect of ionic conditions on DNA helical repeat, effective diameter and free energy of supercoiling. *Nucleic Acids Res*, 25(7):1412–8, 1997.
- [60] V. V. Rybenkov, N. R. Cozzarelli, and A. V. Vologodskii. Probability of DNA knotting and the effective diameter of the dna double helix. *Proc Natl Acad Sci U S A*, 90(11):5307–11, 1993.
- [61] E. G. Yarmola, M. I. Zarudnaya, and S. Lazurkin Yu. Osmotic pressure of DNA solutions and effective diameter of the double helix. *J Biomol Struct Dyn*, 2(5):981–93, 1985.
- [62] J. Huang, T. Schlick, and A. Vologodskii. Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition. *J Mol Biol*, 284(2):287–96, 1998.
- [63] R. Kanaar, A. Klippel, E. Shekhtman, J. M. Dungan, R. Kahmann, and N. R. Cozzarelli. Processive recombination by the phage Mu Gin system: implications for the mechanisms of DNA strand exchange, DNA site alignment, and enhancer action. *Cell*, 62(2):353–66, 1990.
- [64] C. Johansson, M. Kamali-Moghaddam, and L. Sundstrom. Integron integrase binds to bulged hairpin dna. *Nucleic Acids Res*, 32:4033–4043, 2004.
- [65] Dale Rolfson. *Knots and links*. Publish or Perish Press, Berkeley, CA, 1971.
- [66] J. Bath, D. J. Sherratt, and S. D. Colloms. Topology of Xer recombination on catenanes produced by λ integrase. *J Mol Biol*, 289(4):873–83, 1999.
- [67] R. H. Hoess, A. Wierzbicki, and K. Abremski. The role of the *loxP* spacer region in P1 site-specific recombination. *Nucleic Acids Res*, 14(5):2287–300, 1986.

- [68] M.M. Cox. *Mobile DNA*, chapter DNA inversion in the 2 μ m plasmid of *Saccharomyces cerevisiae*, pages 661–670. Am Soc Microbiology, Washington DC, berg, d.e. and m.m. howe, eds edition, 1989.
- [69] S. J. Spengler, A. Stasiak, and N. R. Cozzarelli. The stereostructure of knots and catenanes produced by phage λ integrative recombination: implications for mechanism and DNA structure. *Cell*, 42(1):325–34, 1985.
- [70] N. J. Crisona, R. Kanaar, T. N. Gonzalez, E. L. Zechiedrich, A. Klippel, and N. R. Cozzarelli. Processive recombination by wild-type Gin and an enhancer-independent mutant. Insight into the mechanisms of recombination selectivity and strand exchange. *J Mol Biol*, 243(3):437–57, 1994.
- [71] D. Buck. The hybrid Int-Xer system and its tangle solutions. *Preprint*, 2007.
- [72] K. Abremski, B. Frommer, A. Wierzbicki, and R. H. Hoess. Properties of a mutant Cre protein that alters the topological linkage of recombination products. *J Mol Biol*, 202(1):59–66, 1988.
- [73] H. W. Benjamin, M. M. Matzuk, M. A. Krasnow, and N. R. Cozzarelli. Recombination site selection by Tn3 resolvase: topological tests of a tracking mechanism. *Cell*, 40(1):147–58, 1985.
- [74] C. Ernst and D. W. Sumners. The growth of the number of prime knots. *Math. Proc. Cambridge Philos. Soc.*, 102(2):303–315, 1987.

DEPARTMENT OF MATHEMATICS AND CENTRE FOR BIOINFORMATICS, IMPERIAL COLLEGE LONDON, LONDON ENGLAND SW7 2AZ UK

E-mail address: `d.buck@imperial.ac.uk`

DEPARTMENT OF MATHEMATICS, POMONA COLLEGE, CLAREMONT, CA 91711, USA

E-mail address: `eflapan@pomona.edu`